Structure of a low-population binding intermediate in protein-RNA recognition

Aditi N. Borkar[a], Michael F. Bardaro Jr.[b], Carlo Camilloni[a], Francesco A. Aprile[a], Gabriele Varani[b], and Michele Vendruscolo[a,1]

[a]Department of Chemistry, University of Cambridge, Cambridge CB2 1EW, United Kingdom; and [b]Department of Chemistry, University of Washington, Seattle, WA 98197-1700

Edited by Joseph D. Puglisi, Stanford University School of Medicine, Stanford, CA, and approved April 29, 2016 (received for review October 28, 2015)

The interaction of the HIV-1 protein transactivator of transcription (Tat) and its cognate transactivation response element (TAR) RNA transactivates viral transcription and represents a paradigm for the widespread occurrence of conformational rearrangements in protein-RNA recognition. Although the structures of free and bound forms of TAR are well characterized, the conformations of the intermediates in the binding process are still unknown. By determining the free energy landscape of the complex using NMR residual dipolar couplings in replica-averaged metadynamics simulations, we observe two low-population intermediates. We then rationally design two mutants, one in the protein and another in the RNA, that weaken specific nonnative interactions that stabilize one of the intermediates. By using surface plasmon resonance, we show that these mutations lower the release rate of Tat, as predicted. These results identify the structure of an intermediate for RNA-protein binding and illustrate a general strategy to achieve this goal with high resolution.

RNA structure | NMR spectroscopy | metadynamics | exact RDC restraints | tensor-free method

Essentially all biochemical reactions taking place in living organisms are associated with macromolecular recognition events. A full understanding the molecular mechanisms underlying such events requires the characterization of binding intermediates, which are states that typically have lifetimes of less than a millisecond and may comprise only 5–15% of the conformational space of proteins (1) and nucleic acids (2, 3). Protein-protein and protein-DNA intermediates have recently been characterized at high resolution (4, 5), but despite considerable advances (3, 6–8), high-resolution structures for protein-RNA intermediates have not been reported yet.

To address this problem, we focused on the well-studied process by which HIV, like other lentiviruses, hijacks the host transcription machinery to activate transcription of the viral genome (9–13). In HIV, transactivation (Fig. S1) requires binding of the transactivator of transcription (Tat) protein and the host positive transcription elongation factor b (P-TEFb) complex (11) to the transactivation response element (TAR), a 59-residue RNA stem-loop (Fig. 1 and Fig. S1) with a highly dynamic structure (10, 12, 13). The NMR structures of free TAR (14–16) and of TAR bound to peptide fragments of Tat and to peptide mimetics of Tat in HIV (16–20) and other lentiviruses (21, 22) revealed the conformational properties of TAR in its free and bound states, and demonstrated that this RNA molecule undergoes significant dynamic rearrangements associated with its functions. Although the TAR–Tat complex has become a paradigm for the widespread occurrence of conformational rearrangements and molecular adaptation in protein-RNA recognition, the pathway and intermediates linking the free and bound states of TAR are still unknown.

Results and Discussion

Determination of the Tat-TAR Free Energy Landscape. Following an approach recently described for proteins (4) we identified Tat-TAR binding intermediates from an analysis of the free energy landscape of bound TAR. To implement this method, we constructed the free energy landscape of TAR bound to a cyclic peptide mimetic of Tat (Tat_{pep}) (20) (*Methods*) by performing replica-averaged metadynamics (RAM) simulations (23) using NMR residual dipolar couplings (RDCs) as structural restraints (Fig. S2). In these simulations, RDCs measured for the C8-H8, C2-H2, C5-H5, C1′-H1′, and C4′-H4′ bonds in TAR RNA (24) (*Methods* and Table S1) were incorporated using the recently proposed ϑ-method (25). These simulations exploit the fact that NMR measurements are time- and ensemble-averaged, and it is thus possible to use them for defining the wide range of conformations populated even by highly dynamical RNA systems such as TAR (3, 5, 26–28).

We validated the results of the RAM simulations by back-calculating the values of the restrained and nonrestrained RDC data (Table S2), as well as additional nonrestrained NMR data, including NOEs (Fig. S3) and J-couplings (Fig. S4), from the RAM and unrestrained [molecular dynamics (MD); *Methods*] ensembles and an NMR structure [Protein Data Bank (PDB) ID code 2KDQ] (20) of bound TAR ("static"). In all cases, we found very good agreement between experiments and calculations, whereby the RAM ensemble is consistently able to reproduce the NMR data better compared with both the MD ensemble and the static structure (Table S3).

Free Energy Minima in the Tat-TAR Free Energy Landscape. The free energy landscape that we determined from the RDC data reveals three major free energy minima (Fig. 2*A* and Table S4). State I

Significance

All biochemical reactions in living organisms require molecular recognition events. In particular, the interactions between protein and RNA molecules are crucial in the regulation of gene expression. However, the transient nature of the conformations populated during the recognition process has prevented a detailed characterization of the mechanisms by which these interactions take place. To address this problem, we report a high-resolution structure of an intermediate state in protein-RNA recognition. We determined this structure by using NMR measurements as ensemble-averaged structural restraints in metadynamics simulations, and validated it by performing a structure-based design of two mutants with rationally modified binding rates.

Author contributions: A.N.B., C.C., G.V., and M.V. designed research; A.N.B., M.F.B., C.C., F.A.A., G.V., and M.V. performed research; A.N.B., M.F.B., C.C., F.A.A., G.V., and M.V. contributed new reagents/analytic tools; A.N.B., M.F.B., C.C., F.A.A., G.V., and M.V. analyzed data; and A.N.B., C.C., F.A.A., G.V., and M.V. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

Data deposition: The atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.pdb.org (PDB ID codes 5J0M, 5J1O, and 5J2W).

[1]To whom correspondence should be addressed. Email: mv245@cam.ac.uk.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1521349113/-/DCSupplemental.

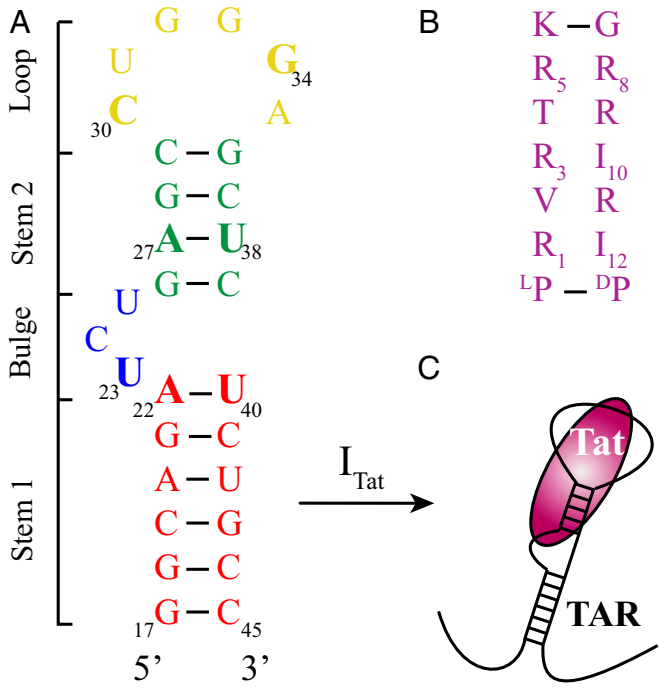

Fig. 1. Schematic illustration of the sequence and secondary structures of TAR and of the Tat mimetic used in this study. Apical region (residues 17–45) of the HIV-1 TAR RNA element (*A*) and Tat_{pep} (*B*) (20). (*C*) Tat binds to the bulge region of TAR, whereas apical loop residues are contacted by the host kinase P-TEFb (11) to form a strongly cooperative ternary complex (50) (Fig. S1). We denote as I_{Tat} the intermediate state on this pathway.

(75% of the sampled conformations) corresponds to the ground state of the complex and is characterized by the presence of essentially all of the TAR-Tat_{pep} native contacts (<0.35 nm) (Fig. 2*B*). States II and III (15% and 7% of the sampled conformations, respectively) correspond to intermediate states of low population. State II is distinct from both the free and bound TAR structures and exhibits a major structural difference in the regions of TAR in contact with Tat_{pep}, compared with the bound state (Fig. 2*B*), consisting of sliding of the peptide away from the apical residues toward stem I of TAR and significant restructuring of its apical loop and bulge. State III is closer to the structure of free TAR than to the structure of bound TAR. In this third conformational minimum, nearly all Tat_{pep} residues lose native contacts with the two helical stems, bulge and residue G34 in the apical loop of TAR. Conversely, U25 in the TAR bulge gains a significant number of nonnative Tat_{pep} contacts (Fig. 2*B*).

Identification of the Tat-TAR Binding Intermediate. To assess whether these intermediates are associated with the transition of TAR between the free and bound forms, we considered whether the global and local characteristics of TAR in these states resemble the characteristics of the free state (Tables S4 and S5), and whether the native contacts between the bulge residues of TAR and the residues of Tat_{pep} that are critical for binding (R3, R5, and I10) are lost (Table S6). With these criteria, our calculations indicate that state III, which is also higher in free energy than state II, has the characteristics of an intermediate (referred to as I_{Tat}; Fig. 3 *B* and *D*) on the binding pathway of Tat and TAR.

To validate state III as a binding intermediate, we used its structure to design mutations rationally to alter the binding process in a predictable manner. To achieve this result, we identified specific interactions that stabilize the intermediate but are absent in the ground state. We compared pairwise native contacts between TAR and Tat_{pep} residues in the I_{Tat} and ground state structures (Fig. 3, Figs. S5–S8, and Table S6), focusing on isolating nonnative interactions around the TAR bulge that stabilize I_{Tat}. The criteria used for selecting suitable candidates for the structure-based design of mutations for validation of I_{Tat} are as follows: (*i*) An interaction should be predominantly present in the intermediate state and absent in the ground state, (*ii*) it should be a unique feature that does not involve multiple residues from TAR or Tat_{pep}, and (*iii*) a mutation that destabilizes it should minimally perturb the free energy landscape of bound TAR while reducing the rate of release of Tat_{pep}. We observed that a nonnative hydrogen bond is stably formed between the O_2 of U25 and the η-amino group of R5 in 20% of I_{Tat} structures and is essentially absent in the ground state (Fig. 3 *C* and *D*). Thus, the specific mutations (Fig. 3*D*) (*i*) 2-thio-uracil at position 25 of TAR (2-thio-U25) and (*ii*) R5K in Tat_{pep} can be expected to destabilize the intermediate state while minimally perturbing the ground state, thereby reducing the rate of release of Tat from TAR.

Validation of the Tat-TAR Binding Intermediate. To validate the structure of I_{Tat} using these mutants, we then used surface plasmon resonance (SPR) experiments to obtain kinetic data for the mutant and the wild-type TAR–Tat_{pep} complexes (Fig. 4). As predicted, we observe a slight decrease in dissociate rate (K_{off}, −17%) for the 2-thio-U25 mutant compared with the wild-type complex (Fig. 4 *B* and *E*), consistent with the weakening of the hydrogen bond induced by the oxygen-to-sulfur substitution (29) in the intermediate, resulting in an overall decrease in K_d (−250%). Also in the R5K mutant, we observed a decrease in K_{off} (−800%) (Fig. 4 *A* and *D*), whose substantial value stems from the fact that through this single mutation, we removed three interactions within the TAR–Tat_{pep} complex: (*i*) the U25:R5 H-bond, (*ii*) the R5:G28 pairing, and (*iii*) the cation-π stacking interaction between R5 and U23 (Fig. 3 *C* and *D*). The latter two interactions are, however, also present in the ground state structures and impart major stability to the TAR–Tat_{pep} complex in general. Thus, the R5K mutation not only destabilizes I_{Tat} but also reduces the overall affinity of the Tat_{pep} for TAR; thus, we also observe a significant increase in K_d (+15-fold) for this mutant compared with the wild-type complex (Fig. 4*A*). Thus, as predicted, the SPR experiments show that when nonnative hydrogen bond interactions gained in I_{Tat} are weakened, the rate of release of Tat from the complex also decreases, thus establishing its relevance as an intermediate between the free and bound forms of TAR induced by peptide binding.

Conclusions

The HIV TAR–Tat interaction has been a subject of major attention in the past two decades, both for understanding the mechanism of transactivation and for development of anti-HIV therapeutics (9–21, 30), and as a paradigm for the mechanism underlying protein-RNA recognition and signaling observed in a wide range of posttranscriptional regulatory processes. Our results reveal the structure of an intermediate in this interaction, illustrating how the use of RDCs as structural restraints in RAM simulations, particularly with further experimental validation through structure-based mutant design, provides a general strategy for obtaining high-resolution structures of low-population intermediates of RNA–protein complexes, which are very challenging for more conventional structure determination or dynamic techniques.

Methods

Bound Structure of TAR. As a starting point for the calculations, and as a reference conformation to analyze the results, we used a previously determined structure of the HIV-1 TAR bound to a 14-residue, cyclic peptidomimetic of the HIV-1 Tat protein (PDB ID code 2KDQ) (20). In the absence of other binding

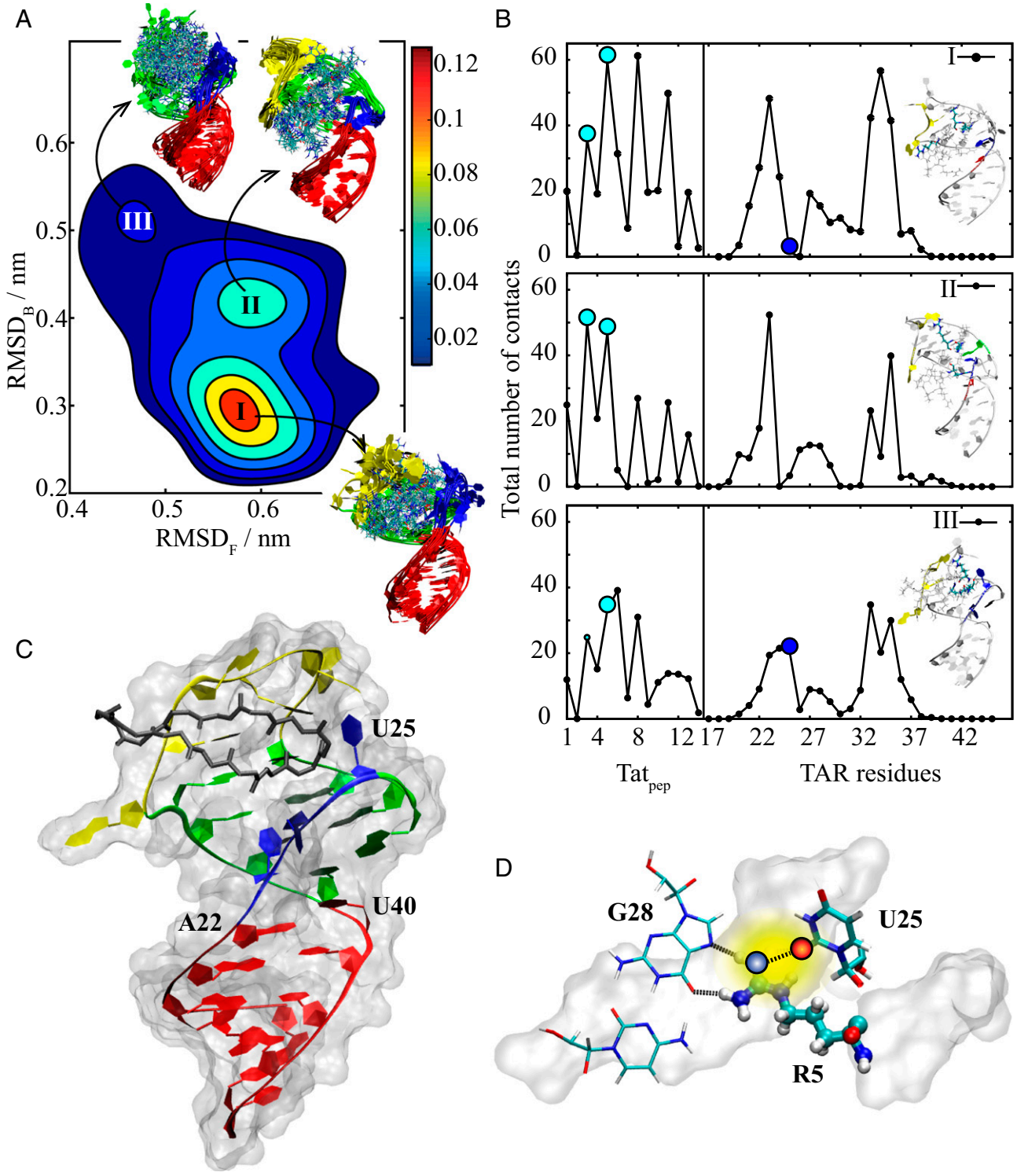

Fig. 2. Free energy landscape of the TAR–Tat complex. (*A*) Probability distribution of the RAM ensemble. The two coordinates used in the plot (which were not used as CVs in the RAM simulations) are the rmsd values ($RMSD_F$ and $RMSD_B$) of the TAR conformations in the RAM ensemble from existing NMR structures of free or bound TAR (14, 20). Three distinct minima are observed: State I is the ground state of bound TAR with about 75% of the total sampled conformations, whereas states II and III are intermediate states with 15% and 7%, respectively, of the sampled conformations. Structures present within these minima are represented as ribbons for TAR and as solid bonds for the peptide. The coloring scheme for TAR ribbon representation is similar to the coloring scheme followed in Fig. 1. The color bar indicates probability densities. (*B*) Total number of contacts (<0.35 nm) made by Tat_{pep} residues (1–14) with TAR nucleotides (17–45), and vice versa. The cyan-filled circles highlight the significant differences in contacts made by R3 and R5 between the ground and excited states of bound TAR, whereas blue-filled circles highlight significant differences in contacts formed by U25. These distinct interactions arise within the three states due to the dynamics of TAR. (*Insets*) Depiction of the location of the top five interactions in the corresponding conformation (ground and intermediate states) of bound TAR. Because state III loses most of the native contacts between TAR and Tat, we identify it as an intermediate (denoted as I_{Tat}) along the pathway of binding of Tat by TAR. As shown in the structure of I_{Tat}, the signature A22:U40 base pair is broken (*C*) and U25 makes specific, nonnative contacts with R5 (*D*). The additional loss of the planar A27:U38:U23 base triplet is denoted as a broken surface below the G28, U25, and R5 interactions.

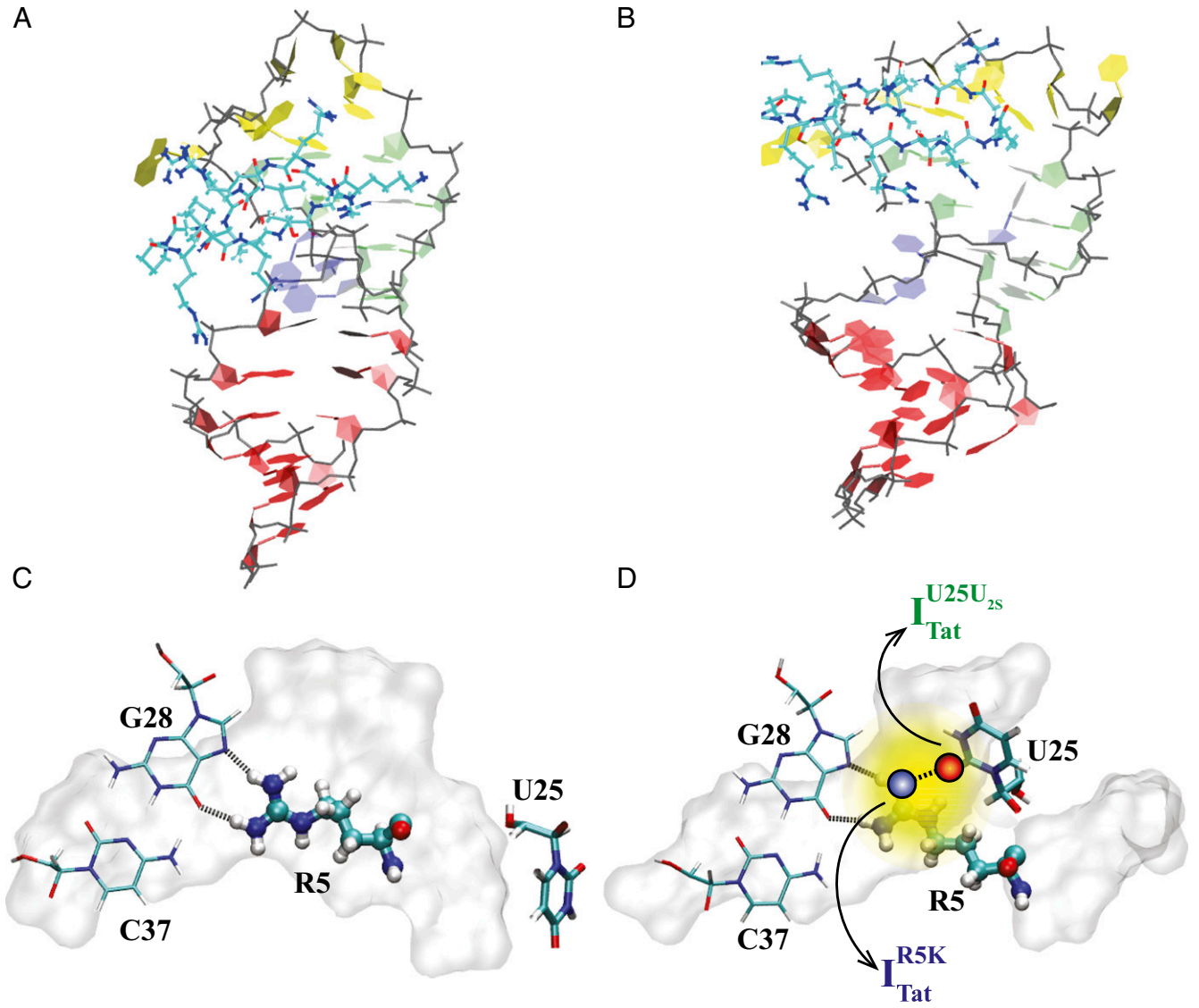

Fig. 3. Structure-based design of mutational variants to validate I_{Tat}. We compared the atomic-resolution structures of the ground state of the TAR–Tat complex (*A*) and of the binding intermediate, I_{Tat} (*B*), to design mutations rationally to alter the TAR-Tat binding kinetics; the color code is the same as in Fig. 1 *A* and *B*. In the ground state, U25 is unpaired and unstacked (*C*), whereas in I_{Tat}, it makes nonnative hydrogen bonds with the η-nitrogen (blue-filled circle) within the guanidium group of R5 through its O_2 function (red-filled circle) (*D*). In these panels, the A27:U38:U23 base triplet is denoted as a solid surface, other TAR residues are illlustrated in stick representations, and R5 is shown as a thick CPK model. Destabilizing this interaction by mutating U25 to 2-thio-U25 ($I_{Tat}^{U25U\text{-}2S}$) or R5 to K5 (I_{Tat}^{R5K}) is predicted to destabilize I_{Tat}. These mutations are expected to diminish the rate of release (K_{off}) of Tat from bound TAR, as can be quantitated by SPR measurements (Fig. 4).

partners, such as the human cyclin T1 and the Cdk9 kinase (11), Tat only makes transient contacts with the TAR RNA and remains unfolded (17). TAR itself only makes a partially ordered structure. Altogether, these properties have prevented the high-resolution structural characterization of the wild-type TAR–Tat complex to date, even if the complex of Tat and P-TEFb was completed (11). Thus, as a model for Tat, we have used the cyclic peptide that binds to TAR with greater specificity than the protein itself (20). Because of the similar characteristics of the bound structures, the conclusions drawn for the binding mechanism of Tat_{pep} are very likely to hold true for the wild-type Tat (20).

Measurement of RDCs. RDCs were measured in multiple alignment media for the bound form of TAR as described previously (24). A summary of the media used and of the number and types of the measured RDCs is provided in Table S1.

RAM Simulations Using the ϑ-Method. Metadynamics is an approach for constructing the free energy landscape of a system by allowing it to escape local minima to explore the conformational space more efficiently (31), thus accelerating the rate of sampling of rare events. In this work, we used it to sample more effectively the free energy landscape corresponding to the force field modified using RDC restraints (discussed below). Metadynamics requires preliminary identification of collective variables (CVs) that describe the phenomenon of interest and are used to construct the free energy landscape of the system. In this work, we carried out RAM simulations (23, 25) to incorporate NMR measurements as structural restraints according to the maximum entropy principle (32–35). RAM simulations were carried out by using the implementation in PLUMED version 2 (36).

Because we have set out to understand mechanistically how TAR transitions between two distinct conformational states, we chose as CVs four structural characteristics of TAR that can distinguish between its free and bound forms. These characteristics are the A22:U40 (17, 37) and C30:G34 (17, 38, 39) base pairs; the U23:A27:U38 base triplet (16, 17, 19, 40, 41); and the extent of base stacking between the bulge residues U23, C24, and U25 (16, 17, 37). The CVs for the base pairs and triplet were implemented in a way that measures the number of hydrogen bonds between the residues involved in the interaction. For example, a value of 0 for the A22:U40 CV

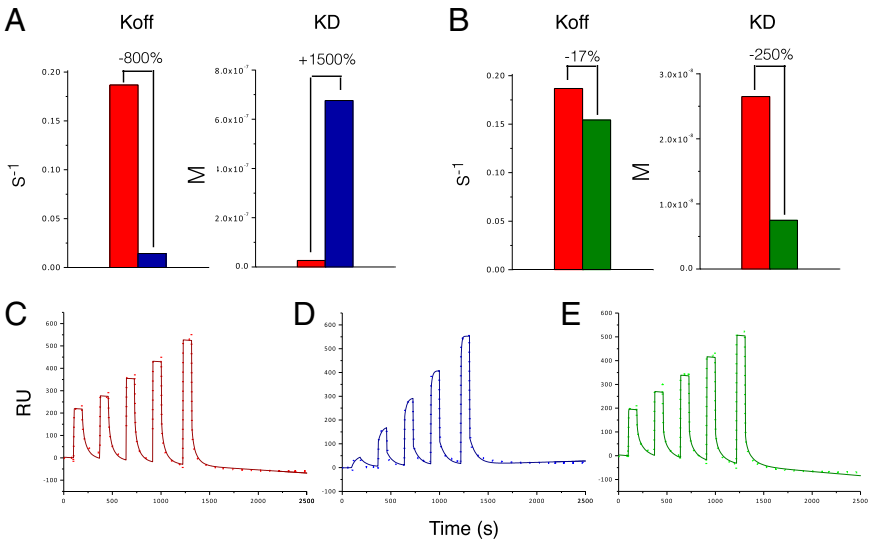

Fig. 4. SPR validation of I_{Tat} using the rationally designed mutants. SPR measurements for validation of I_{Tat}. In I_{Tat}^{R5K}, the K_{off} decreases significantly (by 800%, blue) compared with the wild type (red) (*A*), and in $I_{Tat}^{U25U\text{-}2S}$, it decreases by 17% (green) (*B*). (*C*–*E*) Individual SPR profile fittings for the wild-type I_{Tat}, I_{Tat}^{R5K}, and $I_{Tat}^{U25U\text{-}2S}$. M, molar; RU, resonance units.

signifies that the base pair is completely absent and a value of 2 denotes that the base pair is stably formed. The extent of stacking between the bulge residues was calculated as the distance (as mean square deviation) of the instantaneous conformation in the simulation along the path joining a representative free TAR conformation on one end and the bound form conformation on the other end. These values are set to range from 1 for conformations identical to the free TAR bulge conformation to up to 2 for conformations that are identical to the bound TAR bulge conformation.

For each conformation of TAR sampled during the simulations, RDCs were back-calculated using the ϑ-method (25) (*Supporting Information*) using the implementation in PLUMED version 2 (36). The restraint forces were derived from the agreement between the experimental and calculated RDCs averaged over the ensemble of instantaneous structures from all of the simulated replicas. This approach generates an ensemble of structures according to the maximum entropy principle (32–35).

All of the RDCs measured experimentally were pruned to remove those RDCs with ≤2-Hz bond values and those RDCs that were back-predicted poorly [via the singular value decomposition (SVD) method (42)], which, when removed from the set, improved the prediction of the other RDCs, on a trial set of TAR conformations generated by short restrained and unrestrained MD simulations. After this pruning, we only restrained this set during the simulations and the remaining data were used as "free data" for validation of the simulations. Using Pf1 RDCs for ϑ-restraints is additionally advantageous because this medium is very widely used for aligning both proteins and nucleic acids, but modeling of the Pf1 phage and its electrostatic interactions with RNA involves several approximations that reduce accuracy (43).

MD Simulations. All simulations were performed in the GROMACS 4.5 package (44) using the Amber99bsc0 force field with the χ-parametrization (45). The NMR structure of TAR bound to Tat_{pep} (PDB ID code 2KDQ) (20) was used as the starting conformation for the bound TAR simulations. One model from this structure was placed in a truncated octahedron box with sides 12 Å away from the molecule and solvated with transferable intermolecular potential with 3 points (TIP3P) water (46) molecules. In addition to neutralizing K^+ ions, 100 mM $MgCl_2$ was used. All ions were placed using the genion utility in GROMACS, which randomly replaces water molecules with monoatomic ions. The ion parameters were sourced from the Amber99 force field. After energy minimization with the steepest descent method and then with a low-memory Broyden–Fletcher–Goldfarb–Shanno quasi-Newtonian minimizer, the system was simulated for 50 ps at 200 K without any pressure coupling and with position restraints on both the RNA and peptide. Subsequently, these position restraints were removed, and the temperature of the system was raised to 298.15 K while simulating under constant number of particles, volume, and absolute temperature (NVT) conditions of the system for 100 ps. The macromolecule and its environment (water and ions together) were separately coupled to a Nosé–Hoover thermostat (47) with a coupling constant of 2 ps to maintain the temperature of the system. Next, the system was coupled to the Parinello–Rahman barostat (48) with a coupling constant of 1.5 ps, and eight different simulations were started under constant number of particles, pressure, and absolute temperature (NPT) conditions with different velocity seeds. These temperature and pressure coupling constants were specifically optimized to yield stable simulations with best statistics for control of temperature and pressure of the system (data not reported). During the NPT simulations, the eight trajectories were frequently analyzed for pairwise correlation between the root mean square deviation (rmsd) of the sampled conformations and were terminated when the average correlation between the rmsd values dropped to 0.26 (after ∼5 ns).

Subsequently, RDC restraints were switched on, and the eight NPT trajectories were simulated in parallel as replicas in a single run. First, only the correlation between the experimental and calculated RDCs was restrained using a high force constant to obtain a value of ∼1 (25), and the effective scaling factor was calculated as the inverse of the slope of the calculated vs. experimental RDC values. Thereafter, the Q-factor was restrained to ∼0 using this optimized scaling factor and an optimized force constant value (2,000 kJ/mol) until the autocorrelation of the Q-factor for the ensemble-averaged RDCs of the eight replicas fell to 0 (a further ∼3 ns). Finally, metadynamics were switched on in the bias-exchange mode. Each CV as described above was imposed on two replicas, and the trajectories were simulated until they converged.

For comparison, a RAM simulation using the same setup as described above but without the RDC restraints was also performed separately for bound TAR, resulting in the MD ensemble shown in Figs. S3, S4, and S9.

Test of Convergence of the Simulations. The convergence of the simulations was tested for all simulations using the *sum_hills* utility in PLUMED version 2 (36) and METAGUI software (49). This procedure classified the statistically significant sampled microstates into three basins, which correspond to the three minima observed on the free energy surfaces of the TAR–Tat complex (Fig. 2*A*). The population of each minimum was thus calculated as the fraction of structures clustered into each of these three basins. Simulations were considered to be converged if the history-dependent free energy profiles constructed via the *sum_hills* are similar and the fluctuations between progressive profiles are minimal for the low-energy regions (Fig. S2). Additionally, we continued the simulations until the free energy difference between two small windows on the profiles equilibrated to about 0. This analysis also provided us with an estimate of the equilibration time t_{eq} to input into METAGUI. In METAGUI, simulations were considered to be converged if the differences between free energy profiles during the time intervals $(t_{eq}, t_{eq}+t_{sim}/2)$ and $(t_{eq}+t_{sim}/2, t_{sim})$ do not exceed kT (where t_{sim} is the total simulation time, k is the Boltzmann constant, and T is the temperature in Kelvin at which the simulations were performed). Finally, structures from the converged parts of the simulations were weighted according to their free energies using METAGUI software and used to construct the RAM and MD ensembles.

Whereas the RAM ensemble revealed three distinct states (Fig. 2*A*), the control simulation (MD) carried out using the same protocol but without RDC restraints resulted in the identification of a single free energy minimum (Fig. S9). Similarly, by using two additional sets of RDCs (acrylamide gel and glucopon-hexanol; Table S1), we carried out control RAM simulations to obtain two additional free energy landscapes of the TAR-Tat bound state (Fig. S10). In both cases, we found the ground state to be close to the ground state obtained using Pf1 (Fig. 2*A*), although for these control simulations, the smaller number of RDC restraints (Table S1) in key positions in the TAR molecule (Fig. S11) resulted in intermediate state structures less accurate than those corresponding structure determined in Pf1 (Fig. 2*A*).

ACKNOWLEDGMENTS. A.N.B. was supported by a Dr. Manmohan Singh Scholarship and a Sir Henry Wellcome Postdoctoral Fellowship. M.F.B. and G.V. were supported by grants from the National Science Foundation and NIH.

1. Neudecker P, et al. (2012) Structure of an intermediate state in protein folding and aggregation. *Science* 336(6079):362–366.
2. Al-Hashimi HM (2013) NMR studies of nucleic acid dynamics. *J Magn Reson* 237: 191–204.
3. Dethoff EA, Petzold K, Chugh J, Casiano-Negroni A, Al-Hashimi HM (2012) Visualizing transient low-populated structures of RNA. *Nature* 491(7426):724–728.
4. De Simone A, Aprile FA, Dhulesia A, Dobson CM, Vendruscolo M (2015) Structure of a low-population intermediate state in the release of an enzyme product. *eLife* 4:4.
5. Clore GM, Iwahara J (2009) Theory, practice, and applications of paramagnetic relaxation enhancement for the characterization of transient low-population states of biological macromolecules and their complexes. *Chem Rev* 109(9):4108–4139.
6. Cruz JA, Westhof E (2009) The dynamic landscapes of RNA architecture. *Cell* 136(4): 604–609.
7. Zhao B, Zhang Q (2015) Characterizing excited conformational states of RNA by NMR spectroscopy. *Curr Opin Struct Biol* 30:134–146.
8. Krepl M, et al. (2015) Can we execute stable microsecond-scale atomistic simulations of protein–RNA complexes? *J Chem Theory Comput* 11(3):1220–1243.
9. Peterlin BM, Price DH (2006) Controlling the elongation phase of transcription with P-TEFb. *Mol Cell* 23(3):297–305.
10. Bardaro MF, Jr, Shajani Z, Patora-Komisarska K, Robinson JA, Varani G (2009) How binding of small molecule and peptide ligands to HIV-1 TAR alters the RNA motional landscape. *Nucleic Acids Res* 37(5):1529–1540.
11. Tahirov TH, et al. (2010) Crystal structure of HIV-1 Tat complexed with human P-TEFb. *Nature* 465(7299):747–751.
12. Zhang Q, Sun X, Watt ED, Al-Hashimi HM (2006) Resolving the motional modes that code for RNA adaptation. *Science* 311(5761):653–656.
13. Zhang Q, Stelzer AC, Fisher CK, Al-Hashimi HM (2007) Visualizing spatially correlated dynamics that directs RNA conformational transitions. *Nature* 450(7173):1263–1267.
14. Aboul-ela F, Karn J, Varani G (1996) Structure of HIV-1 TAR RNA in the absence of ligands reveals a novel conformation of the trinucleotide bulge. *Nucleic Acids Res* 24(20):3974–3981.
15. Ippolito JA, Steitz TA (1998) A 1.3-A resolution crystal structure of the HIV-1 transactivation response region RNA stem reveals a metal ion-dependent bulge conformation. *Proc Natl Acad Sci USA* 95(17):9819–9824.
16. Long KS, Crothers DM (1999) Characterization of the solution conformations of unbound and Tat peptide-bound forms of HIV-1 TAR RNA. *Biochemistry* 38(31): 10059–10069.
17. Aboul-ela F, Karn J, Varani G (1995) The structure of the human immunodeficiency virus type-1 TAR RNA reveals principles of RNA recognition by Tat protein. *J Mol Biol* 253(2):313–332.
18. Hamy F, et al. (1997) An inhibitor of the Tat/TAR RNA interaction that effectively suppresses HIV-1 replication. *Proc Natl Acad Sci USA* 94(8):3548–3553.
19. Puglisi JD, Chen L, Frankel AD, Williamson JR (1993) Role of RNA structure in arginine recognition of TAR RNA. *Proc Natl Acad Sci USA* 90(8):3680–3684.
20. Davidson A, et al. (2009) Simultaneous recognition of HIV-1 TAR RNA bulge and loop sequences by cyclic peptide mimics of Tat protein. *Proc Natl Acad Sci USA* 106(29): 11931–11936.
21. Puglisi JD, Chen L, Blanchard S, Frankel AD (1995) Solution structure of a bovine immunodeficiency virus Tat-TAR peptide-RNA complex. *Science* 270(5239):1200–1203.
22. Anand K, Schulte A, Vogel-Bachmayr K, Scheffzek K, Geyer M (2008) Structural insights into the cyclin T1-Tat-TAR RNA transcription activation complex from EIAV. *Nat Struct Mol Biol* 15(12):1287–1292.
23. Camilloni C, Cavalli A, Vendruscolo M (2013) Replica-Averaged Metadynamics. *J Chem Theory Comput* 9(12):5610–5617.
24. Bardaro MF, Jr, Varani G (2012) Independent alignment of RNA for dynamic studies using residual dipolar couplings. *J Biomol NMR* 54(1):69–80.
25. Camilloni C, Vendruscolo M (2015) A tensor-free method for the structural and dynamical refinement of proteins using residual dipolar couplings. *J Phys Chem B* 119(3): 653–661.
26. Lindorff-Larsen K, Best RB, Depristo MA, Dobson CM, Vendruscolo M (2005) Simultaneous determination of protein structure and dynamics. *Nature* 433(7022):128–132.
27. Sekhar A, Kay LE (2013) NMR paves the way for atomic level descriptions of sparsely populated, transiently formed biomolecular conformers. *Proc Natl Acad Sci USA* 110(32):12867–12874.
28. Borkar AN, De Simone A, Montalvao RW, Vendruscolo M (2013) A method of determining RNA conformational ensembles using structure-based calculations of residual dipolar couplings. *J Chem Phys* 138(21):215103.
29. Camilloni C, et al. (2014) Cyclophilin A catalyzes proline isomerization by an electrostatic handle mechanism. *Proc Natl Acad Sci USA* 111(28):10203–10208.
30. Al-Hashimi HM, et al. (2002) Concerted motions in HIV-1 TAR RNA may allow access to bound state conformations: RNA dynamics from NMR residual dipolar couplings. *J Mol Biol* 315(2):95–102.
31. Laio A, Parrinello M (2002) Escaping free-energy minima. *Proc Natl Acad Sci USA* 99(20):12562–12566.
32. Pitera JW, Chodera JD (2012) On the use of experimental observations to bias simulated ensembles. *J Chem Theory Comput* 8(10):3445–3451.
33. Cavalli A, Camilloni C, Vendruscolo M (2013) Molecular dynamics simulations with replica-averaged structural restraints generate structural ensembles according to the maximum entropy principle. *J Chem Phys* 138(9):094112.
34. Roux B, Weare J (2013) On the statistical equivalence of restrained-ensemble simulations with the maximum entropy method. *J Chem Phys* 138(8):084107.
35. Boomsma W, Ferkinghoff-Borg J, Lindorff-Larsen K (2014) Combining experiments and simulations using the maximum entropy principle. *PLOS Comput Biol* 10(2): e1003406.
36. Tribello GA, Bonomi M, Branduardi D, Camilloni C, Bussi G (2014) Plumed 2: New feathers for an old bird. *Comput Phys Commun* 185:604–613.
37. Pitt SW, Majumdar A, Serganov A, Patel DJ, Al-Hashimi HM (2004) Argininamide binding arrests global motions in HIV-1 TAR RNA: Comparison with Mg2+-induced conformational stabilization. *J Mol Biol* 338(1):7–16.
38. Richter S, Cao H, Rana TM (2002) Specific HIV-1 TAR RNA loop sequence and functional groups are required for human cyclin T1-Tat-TAR ternary complex formation. *Biochemistry* 41(20):6391–6397.
39. Dethoff EA, et al. (2008) Characterizing complex dynamics in the transactivation response element apical loop and motional correlations with the bulge by NMR, molecular dynamics, and mutagenesis. *Biophys J* 95(8):3906–3915.
40. Nifosì R, Reyes CM, Kollman PA (2000) Molecular dynamics studies of the HIV-1 TAR and its complex with argininamide. *Nucleic Acids Res* 28(24):4944–4955.
41. Tao J, Chen L, Frankel AD (1997) Dissection of the proposed base triple in human immunodeficiency virus TAR RNA indicates the importance of the Hoogsteen interaction. *Biochemistry* 36(12):3491–3495.
42. Losonczi JA, Andrec M, Fischer MW, Prestegard JH (1999) Order matrix analysis of residual dipolar couplings using singular value decomposition. *J Magn Reson* 138(2):334–342.
43. Montalvao RW, De Simone A, Vendruscolo M (2012) Determination of structural fluctuations of proteins from structure-based calculations of residual dipolar couplings. *J Biomol NMR* 53(4):281–292.
44. Hess B, Kutzner C, van der Spoel D, Lindahl E (2008) Gromacs 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J Chem Theory Comput* 4(3):435–447.
45. Banas P, et al. (2010) Performance of molecular mechanics force fields for RNA simulations: Stability of UUCG and GNRA hairpins. *J Chem Theory Comput* 6(12):3836–3849.
46. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) Comparison of simple potential functions for simulating liquid water. *J Chem Phys* 79:926–935.
47. Hoover WG (1985) Canonical dynamics: Equilibrium phase-space distributions. *Phys Rev A Gen Phys* 31(3):1695–1697.
48. Parrinello M, Rahman A (1981) Polymorphic transitions in single crystals: A new molecular dynamics method. *J Appl Phys* 52:7182–7190.
49. Biarnes X, Pietrucci F, Marinelli F, Laio A (2012) METAGUI. A VMD interface for analyzing metadynamics and molecular dynamics simulations. *Comput Phys Commun* 183:203–211.
50. Camilloni C, Vendruscolo M (2014) Statistical mechanics of the denatured state of a protein using replica-averaged metadynamics. *J Am Chem Soc* 136(25):8982–8991.
51. Laio A, Gervasio FL (2008) Metadynamics: A method to simulate rare events and reconstruct the free energy in biophysics, chemistry and material science. *Rep Prog Phys* 71:126601.
52. Piana S, Laio A (2007) A bias-exchange approach to protein folding. *J Phys Chem B* 111(17):4553–4559.
53. Bax A (2003) Weak alignment offers new NMR opportunities to study protein structure and dynamics. *Protein Sci* 12(1):1–16.
54. Tolman JR, Flanagan JM, Kennedy MA, Prestegard JH (1995) Nuclear magnetic dipole interactions in field-oriented proteins: Information for structure determination in solution. *Proc Natl Acad Sci USA* 92(20):9279–9283.
55. Wirz LN, Allison JR (2015) Comment on "a tensor-free method for the structural and dynamic refinement of proteins using residual dipolar couplings". *J Phys Chem B* 119(25):8223–8224, discussion 8225–8226.
56. Camilloni C, Vendruscolo M (2015) Reply to "comment on 'a tensor-free method for the structural and dynamic refinement of proteins using residual dipolar couplings'". *J Phys Chem B* 119:8225–8226.
57. Olsson S, Ekonomiuk D, Sgrignani J, Cavalli A (2015) Molecular dynamics of biomolecules through direct analysis of dipolar couplings. *J Am Chem Soc* 137(19): 6270–6278.
58. Karn J (1999) Tackling Tat. *J Mol Biol* 293(2):235–254.
59. Salmon L, Bascom G, Andricioaei I, Al-Hashimi HM (2013) A general method for constructing atomic-resolution RNA ensembles using NMR residual dipolar couplings: The basis for interhelical motions revealed. *J Am Chem Soc* 135(14):5457–5466.